

Developmental Change in the Acuity of the “Number Sense”: The Approximate Number System in 3-, 4-, 5-, and 6-Year-Olds and Adults

Justin Halberda and Lisa Feigenson
Johns Hopkins University

Behavioral, neuropsychological, and brain imaging research points to a dedicated system for processing number that is shared across development and across species. This foundational Approximate Number System (ANS) operates over multiple modalities, forming representations of the number of objects, sounds, or events in a scene. This system is imprecise and hence differs from exact counting. Evidence suggests that the resolution of the ANS, as specified by a Weber fraction, increases with age such that adults can discriminate numerosities that infants cannot. However, the Weber fraction has yet to be determined for participants of any age between 9 months and adulthood, leaving its developmental trajectory unclear. Here we identify the Weber fraction of the ANS in 3-, 4-, 5-, and 6-year-old children and in adults. We show that the resolution of this system continues to increase throughout childhood, with adultlike levels of acuity attained surprisingly late in development.

Keywords: analog magnitudes, mathematics, estimation, numerical, intraparietal sulcus, approximation

The ability to nonverbally represent number is shared across species and across development (for reviews see Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004). The foundational Approximate Number System (ANS) that underlies this ability produces abstract number representations (Barth, Kanwisher, & Spelke, 2003) that support arithmetic computation across the life span (Barth et al., 2003; Barth et al., 2006; McCrink & Wynn, 2004). The ANS is activated when adults perform symbolic number tasks (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003; Piazza, Pinel, LeBihan, & Dehaene, 2007) and may even provide a foundation for more sophisticated mathematics (Gilmore, McCarthy, & Spelke, 2007). However, the ANS represents number only approximately (Dehaene, 1997; Gallistel & Gelman, 2000), and the imprecision of its numerical representations is radically greater in infants than in adults (e.g., Pica, Lemer, Izard, & Dehaene, 2004; Xu & Spelke, 2000).

Given that the ANS is thought to play an important role in math learning (Booth & Siegler, 2006; Jordan, Kaplan, Locuniak, & Ramineni, 2007), it is surprising that no research to date has explored the full developmental trajectory of its representational acuity. Here we tested ANS acuity in 3-, 4-, 5-, and 6-year-old children and in adults, using psychophysical modeling to deter-

mine the finest numerical discriminations possible at each age. Our findings reveal that the ANS does not attain full acuity until quite late in development, long after children have begun formal instruction in mathematics.

The ANS differs from counting in that it produces inexact number representations (Gallistel & Gelman, 1992; but see Zorzi & Butterworth, 1999). A hallmark of the ANS is that the imprecision of its representations grows with the target numerosity, such that the ability to nonverbally discriminate two quantities depends on their ratio (Moyer & Landauer, 1967). This ratio dependence is observed when adults estimate numbers of items (Halberda, Sires, & Feigenson, 2006; Whalen, Gallistel, & Gelman, 1999), produce target numbers of actions (Cordes, Gelman, & Gallistel, 2001; Whalen et al., 1999), judge the more numerous of two arrays (Barth et al., 2003), and estimate the results of arithmetic events (Pica et al., 2004). Because of the inexactness of ANS representations, two quantities cannot be distinguished when the distance between them is too small. The finest numerical ratio that adults can consistently discriminate has been identified as 7:8. This limit can also be described as a Weber fraction that measures the smallest numerical change to a stimulus that can be reliably detected. The Weber fraction is equal to the difference between the two numbers divided by the smaller number; for example, $7:8 \rightarrow (8 - 7)/7 = .14$. When asked to indicate the more numerous of two simultaneously presented arrays containing 20–80 dots, French adults' Weber fraction is .12 and Amazonian adults' Weber fraction is .17; thus on average these adults could discriminate ratios differing by about 7:8 (Pica et al., 2004).

Ratio-dependent numerical performance also reveals that preverbal infants use the ANS, albeit with drastically less acuity than adults. Six-month old infants discriminate arrays of 4 versus 8, 8 versus 16, and 16 versus 32 dots, all of which instantiate a 1:2 ratio, but fail to discriminate 8 versus 12 and 16 versus 24, which instantiate a 2:3 ratio (Xu, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). The same pattern obtains in audition: 6-month-olds discriminate

Justin Halberda and Lisa Feigenson, Department of Psychological and Brain Sciences, Johns Hopkins University.

Justin Halberda and Lisa Feigenson functioned as Co-Principal Investigators on this project. We thank Roy Goodman and Fred Heberle for help with modeling and Allison Wessel for help with testing participants.

Guidelines for testing human research subjects were followed as certified by the Johns Hopkins University Institutional Review Boards. Participants' rights were protected throughout.

Correspondence concerning this article should be addressed to Justin Halberda, Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218. E-mail: halberda@jhu.edu

4 versus 8 and 8 versus 16 but not 4 versus 6 or 8 versus 12 tones (Lipton & Spelke, 2003, 2004). Infants' discrimination threshold changes with development: 9-month-olds succeed with the 2:3 ratios with which 6-month-olds fail, in both the auditory and visual domains (Lipton & Spelke, 2003; Wood & Spelke, 2005; Xu & Arriaga, 2007). The identical increase in acuity across sensory modalities suggests that it is the ANS itself, and not the visual or auditory perception systems, that is improving.

The aforementioned data show that ANS acuity increases over the life span, with participants exhibiting a Weber fraction of 1.0 at 6 months, 0.5 at 9 months, and 0.14 in adulthood. However, these studies leave a large gap in our understanding of the development of nonverbal enumeration. It is known that preschool- and early school-age children (the ages during which formal instruction in mathematics typically begins) show performance controlled by numerical ratio, making more errors with close than with distant numerical comparisons (Barth, LaMont, Lipton, & Spelke, 2005; Huntley-Fenner, 2001; Huntley-Fenner & Cannon, 2000; Starkey & Cooper, 1995; Temple & Posner, 1998). Estimates of numerical acuity in 5-, 6-, and 8-year-olds using a numerical bisection task suggest Weber fraction values in the vicinity of .26, intermediate between the acuity of adults and that of infants (Droit-Volet, Clément, & Fayol, 2003; Jordan & Brannon, 2006). An analysis of children's reaction times to determine the larger of two Arabic numbers is also consistent with the proposal that the ANS affects performance such that numerically near digits take longer to discriminate (Sekuler & Mierkiewicz, 1977). This effect changes with age throughout the school years, suggesting the possibility of developmental changes in ANS acuity (Sekuler & Mierkiewicz, 1977).

Yet, no study to date has identified the finest numerical discriminations children can make, as none has used psychophysical methods similar to those used with adults to identify children's Weber fraction. For this reason, the developmental trajectory of changes in ANS acuity between 9 months and adulthood remains undescribed. For example, it is not known whether numerical acuity rapidly asymptotes during the first year of life (similar to stereoacuity; Held, Birch, & Gwiazda, 1980), gradually increases throughout early childhood (similar to executive function; Diamond, 2002; Zelazo, Craik, & Booth, 2004), or shows a discontinuous change when children master verbal counting at around age 4 (similar to a vocabulary spurt; Carey, 1978; Goldfield & Reznick, 1990). Furthermore, changes in ANS acuity have implications for formal instruction in mathematics. Many math curricula aim to tap children's intuitions regarding "possible" or "impossible" solutions to quantitative problems, encouraging children to estimate numerical magnitudes before arriving at an exact answer (Johnson, 1979; Levin, 1981). Given the widespread nature of such teaching tools, it is surprising that children's ANS acuity during the early years of mathematics instruction has not been determined.

We presented 3-, 4-, 5-, and 6-year-old children and adults with a nonverbal number discrimination task that did not permit counting and that included controls for continuous variables that often correlate with number. Varying the numerical ratio between stimulus arrays allowed us to determine the Weber fraction of the ANS for each age group and thereby to describe the function by which the ANS reaches the adult state of representational precision. This cross-sectional approach allowed us to determine whether numerical acuity continues to develop over the ages tested. However, it is important to note that this type of cross-sectional design might

obscure possible distinctions between a group trend and an individual child's developmental trajectory (e.g., individual children might show discontinuous changes in numerical acuity, which would be revealed only in a longitudinal analysis of changes in individual performance over time). Our present goal was to determine whether ANS acuity continues to develop during the early school years and to estimate the group trend in this developmental trajectory.

Method

Participants

We tested five age groups with 16 participants per group: 3-, 4-, 5-, and 6-year-olds and adults. Summaries of mean age, age range, and sex appear in Table 1. The adults were undergraduate and graduate students at the Johns Hopkins University. The children were recruited via mail and telephone, were mostly Caucasian, mostly of middle-socioeconomic status, and mostly with parents with some post-high school education. Each child was tested only once. Thirty-one additional children were tested but not included in the final analysis due to parental interference (2), failure to complete the task (21), or equipment failure (8).

Materials

On each trial of the numerical discrimination task, two arrays of between 1 and 14 items appeared side-by-side on a large video screen (see Figure 1). Items varied in size but were otherwise identical, and on each trial they were randomly chosen from 46 possible images of familiar objects (Table 2). Each array appeared within a background frame demarcating "Big Bird's Xs" on the left side of the screen and "Grover's Xs" on the right.

On each trial, participants heard a recorded voice prompting them to select the greater of the two quantities (e.g., "Who has more [pieces of pizza]?"). Labels and the carrier phrase "Who has more" were recorded by a female native English speaker in a child-friendly voice and were played over a centralized computer speaker.

Procedure

Participants sat at a table approximately 150 cm from the screen (viewing area 120 cm × 90 cm). The average object subtended 2° of visual angle from this distance. The study began with the recorded female voice saying, "Let's play a game," followed by four practice trials. On practice trials the computer first displayed

Table 1
Description of Participants

Group	Mean age	Age range	Sex	
			Male	Female
3 years	3.75	3.33–3.92	10	6
4 years	4.33	4.00–4.92	10	6
5 years	5.41	5.00–5.75	9	7
6 years	6.13	6.00–8.25	10	6
Adults	20.18	18.75–32.25	6	10

Big Bird’s items for 2,000 ms accompanied by the labeling phrase, “Here are Big Bird’s [pieces of pizza].” Next the computer displayed Grover’s items for 2,000 ms accompanied by the labeling phrase, “Here are Grover’s [pieces of pizza].” Finally, both arrays appeared simultaneously for 2,000 ms accompanied by the carrier phrase, “Who has more [pieces of pizza]?” The carrier portion of the phrase began before the items appeared; label onset was synchronized to the items’ visual onset. Participants indicated which character had the greater number of items via a color-coded keyboard, pressing a red key to begin each trial, a yellow key to indicate that Big Bird had more, and a blue key to indicate that Grover had more.

The experimenter and, for children, a caregiver accompanied participants into the testing room and sat approximately 100 cm behind them so as not to influence performance. To maintain participants’ motivation, the computer provided auditory feedback on every trial (e.g., “That’s right!” for correct responses; “Oh, that’s not right” for incorrect responses). Item type and item position within the background frames were free to vary randomly throughout practice and test trials. The 4 practice trials were followed by 66 test trials, which were identical to the simultaneous portion of the practice trials.

Displays were controlled either for average item size (area correlated trials) or summed continuous extent (area anticorrelated trials). For each ratio presented, on half of the trials the larger numerosity had more total surface area (area correlated trials), and on the other half the smaller numerosity had more total surface area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total surface area, two dimensions of continuous extent to which infants have shown sensitivity (Clearfield & Mix, 1999; Feigenson, Carey, & Spelke, 2002). These trials equated the total summed horizontal and total summed vertical extent of the items in Big Bird’s and Grover’s arrays. This procedure equated perimeter and anticorrelated surface area such that the array with the smaller number of items had more total surface area. For example, on an area anticorrelated trial on which Big Bird had 6 items and Grover had 12, Grover would be the numerically correct choice by a ratio of 2 ($12/6 = 2$), but Big Bird would be the correct choice in total area by an equal and opposite ratio ($20,943 \text{ pixels}/1,0471 \text{ pixels} = 2$). Because the difference in total surface area varied with the numerical ratio of the two arrays (as summarized in Table 3), the easier the numerical ratio was to discriminate, the more wrong surface area became. This, combined with consistent feedback, was included to discourage the use of continuous variables as a cue. Finally, on both area correlated and area anticorrelated trials, individual item size varied to ensure that items in the less numer-

Table 2
Objects Used in Experiment 1, Reported According to the Label Used

Bagels	Chairs	Hats	Shovels
Balls	Cherries	Hearts	Squirrels
Bananas	Cookies	Keys	Strawberries
Bicycles	Cows	Ladybugs	Teddy Bears
Blocks	Crayons	Leaves	Tractors
Boots	Cupcakes	Pieces of Pizza	Trains
Bottles	Doggies	Pigs	Trucks
Bows	Dollies	Rings	Umbrellas
Buckets	Donkeys	Sailboats	Violins
Bunnies	Elephants	Sandwiches	Wagons
Butterflies	Flags	School Busses	
Cars	Hammers	Seashells	

ous array were not all larger than those in the more numerous array (see Figure 1).¹

Unequal numbers of trials were presented from each ratio bin in order to focus on more difficult ratios (ratio bin 1:2 = 2 trials, 2:3 = 2, 3:4 = 2, 4:5 = 2, 5:6 = 10, 6:7 = 10, 7:8 = 10, 8:9 = 14, 9:10 = 14). These numbers were initially chosen for adult participants and so focus more heavily on difficult ratios (e.g., 5:6–9:10). To ensure consistency, we used the same ratio distribution for all ages.

Display time was adjusted for each age group through pilot testing and was chosen to be long enough to allow participants to view both arrays but short enough to prevent serial counting.² Display time was 2,500 ms for 3-year-olds, 1,200 ms for 4-, 5-, and 6-year-olds, and 750 ms for adults. The winning side (Big Bird or Grover), ratio presented, trial type (area correlated, area anticorrelated), item type, and absolute number of items presented varied randomly across trials.

Results

To examine whether performance differed across age groups or depended on numerical ratio, we entered each participant’s percentage correct for each ratio bin (1:2, 2:3, etc.) into a 5 (age group) × 2 (sex) × 2 (trial type: area correlated, area anticorrelated) × 9 (ratio)

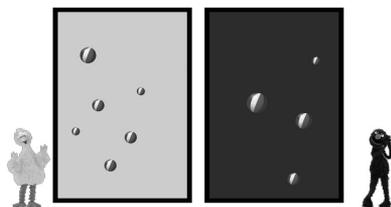


Figure 1. Sample trial from the Who Has More task.

¹ Our method for anticorrelating area and equating the total summed perimeter of the items in Big Bird’s and Grover’s sets relies on the geometric fact that the areas of similar polygons are to each other as the squares of any two corresponding segments. This fact allows us to rely on the horizontal and vertical extent of the image files rather than measuring each minute curve in the perimeter of each picture. Because any curved area can be estimated by a series of connected polygons and because the images used within each trial were scaled variants of a single image, the total summed perimeter was precisely equated on each trial and the ratio of areas was exactly opposite the ratio of the number of items.

² Children’s verbal counting ability was assessed for the purposes of a separate study using Wynn’s Give A Number task (Wynn, 1992) and revealed that none of the 3-year-old children understood the Cardinal Word Principle (Gelman & Gallistel, 1978) and so could not have used counting to determine the numerosities of the sets irrespective of display time.

Table 3
Controls for Continuous Extent on Area Anticorrelated Trials

Whole-number ratio	Numeric ratio (N2/N1)	Total perimeter	Ratio of total area (N1/N2)
1:2	2	Equated	2
2:3	1.5	Equated	1.5
3:4	1.33	Equated	1.33
4:5	1.25	Equated	1.25
5:6	1.2	Equated	1.2
6:7	1.16	Equated	1.16
7:8	1.14	Equated	1.14
8:9	1.13	Equated	1.13
9:10	1.11	Equated	1.11

repeated measures analysis of variance.³ These data are presented in Figure 2, which plots percentage correct for each age group as a function of ratio (numerosity of larger set/numerosity of smaller set). There was a significant age group effect, with participants performing better with increasing age, $F(4, 70) = 28.891, p < .001$; a significant ratio effect, with participants performing better with increasing ratio, $F(8, 560) = 20.199, p < .001$; and a marginally significant effect of sex, with girls and women performing slightly better than boys and men overall, $F(1, 70) = 3.122, p = .082$.

Regarding the effect of total area on participants' numerical discriminations, we observed a significant trial type effect, $F(1, 70) = 6.138, p < .05$. We investigated this effect via planned t tests comparing each age group's percentage correct on area correlated and area anticorrelated trials with chance (50%). Collapsing across all ratios revealed that all age groups performed significantly above chance on both trial types, as summarized in Table 4, but performance was slightly better on area correlated trials. All age groups based their responses on number, not area, though area had some effect on judgments.

We also observed a significant Trial Type \times Ratio interaction, as the difference in performance on area correlated and area anticorrelated trials was larger for easier ratios than for harder ratios, $F(8, 560) = 2.62, p < .01$. A significant Age Group \times Ratio \times Trial Type interaction revealed that older children and adults showed less differentiation in performance between area correlated and area anticorrelated trials as a function of ratio than did younger children. In particular 3-, 4-, and 5-year-olds performed better on area correlated trials than area anticorrelated trials with 1:2 ratio comparisons, whereas 6-year-olds and adults showed little or no difference in performance with this comparison ratio, $F(32, 560) = 1.575, p < .05$. Any effect of area was minor in the experiment, however, as basing answers on area would have resulted in performance that was significantly below chance, and this was never observed.

Participants' performance varied as a function of ratio (see Figure 2). If participants were using the ANS to determine the more numerous array, then percentage correct as a function of ratio (collapsed across trial type) should be well fit by a computational model of the ANS. Pica et al., (2004) examined performance on a task similar to ours in adults and children from both a developed country (France) and an indigenous culture (the Mundurucu of Amazonia) whose language lacks exact large-number words such as "seven." Pica and colleagues found that both groups' perfor-

mance was well fit by a psychophysics model (Green & Swets, 1966; Moyer & Bayer, 1976) that models ANS representations as "noisy" Gaussian random variables and numerical discrimination as the subtraction of the two Gaussian random variables that represent the numerosities of the two arrays.⁴ This model has a single free parameter, the Weber fraction (w), which determines the increase in percentage correct with increasing ratio. We rely on this same psychophysics model, which has been argued to be a parsimonious, psychologically plausible model of numerical performance (Pica et al., 2004). This model has received further support from neuronal data of monkeys performing a numerical discrimination task (Nieder & Miller, 2004).

As has been observed in previous numerical discrimination tasks (Pica et al., 2004), participants' performance can fail to reach 100% correct because of a tendency to guess randomly on some trials. As seen in Figure 2, this was the case for our 3-, 4-, and 5-year-old children. The simplest way to account for this tendency computationally is to include a parameter that is a constant probability of guessing randomly on any particular trial (for additional discussion of this approach see the supporting online materials of Pica et al., 2004). This parameter lowers the model's asymptotic performance while retaining an accurate estimate of the Weber fraction. We included this parameter in our model as follows: where p_{guess} is the probability of guessing randomly, p_{error} is the probability of being incorrect given the model, and chance is .5 multiplied by 100 to return a percentage.

$$\text{percent correct} = [(1 - p_{\text{guess}})(1 - p_{\text{error}}) + p_{\text{guess}} * .5] * 100 \quad (1)$$

As seen in Figures 3a and 3b and Table 5, the psychophysics model provides an accurate fit to our data from 6-year-olds and adults (adult R^2 value = .93). Performance for 3-, 4-, and 5-year-olds deviates from this model, however, because younger children were at chance for more difficult ratios. Figures 3d and 3e show that data from 3- and 4-year-olds exhibit a sigmoidal shape. The likely cause is that the children ceased to rely on ANS representations for the hardest ratios and instead guessed randomly (e.g., 3-year-olds asymptote at 53% for comparisons of ratio 1.25 and lower). This pattern might also simply be a failure to discriminate the numerosities involved on these more difficult ratio comparisons, but performance well above the least squares fit psychophysics model for easier ratios near the elbow of the function (e.g., ratios 1.25, 1.33, and 1.5 for 4-year-olds) suggests that the failures at more difficult ratios are aberrant and a result of "giving up" on

³ Results similar to those reported here were obtained from analyses of reaction time, which was also recorded, and these data show no evidence of speed-accuracy trade-off differences across ages (e.g., 3-year-olds were both slower and less accurate than 4-year-olds). For purposes of psychophysical modeling, we focus our analyses on percentage correct.

⁴ In the psychophysics model, each numerosity is represented as a Gaussian random variable (i.e., $\times 2$ and $\times 1$) with means n_2 and n_1 and standard deviations equal to the Weber fraction (w) $\times n$. Subtracting the Gaussian for the smaller set from that for the larger set returns a new Gaussian that has a mean of $n_2 - n_1$ and a standard deviation of $w\sqrt{n_1^2 + n_2^2}$ (simply the difference of two Gaussian random variables). Percentage correct is then equal to 1 - error rate, where error rate is defined as the area under the tail of the resulting Gaussian curve computed as follows:

$$= \frac{1}{2} \operatorname{erfc} \left(\frac{|n_1 - n_2|}{\sqrt{2}w\sqrt{n_1^2 + n_2^2}} \right)$$

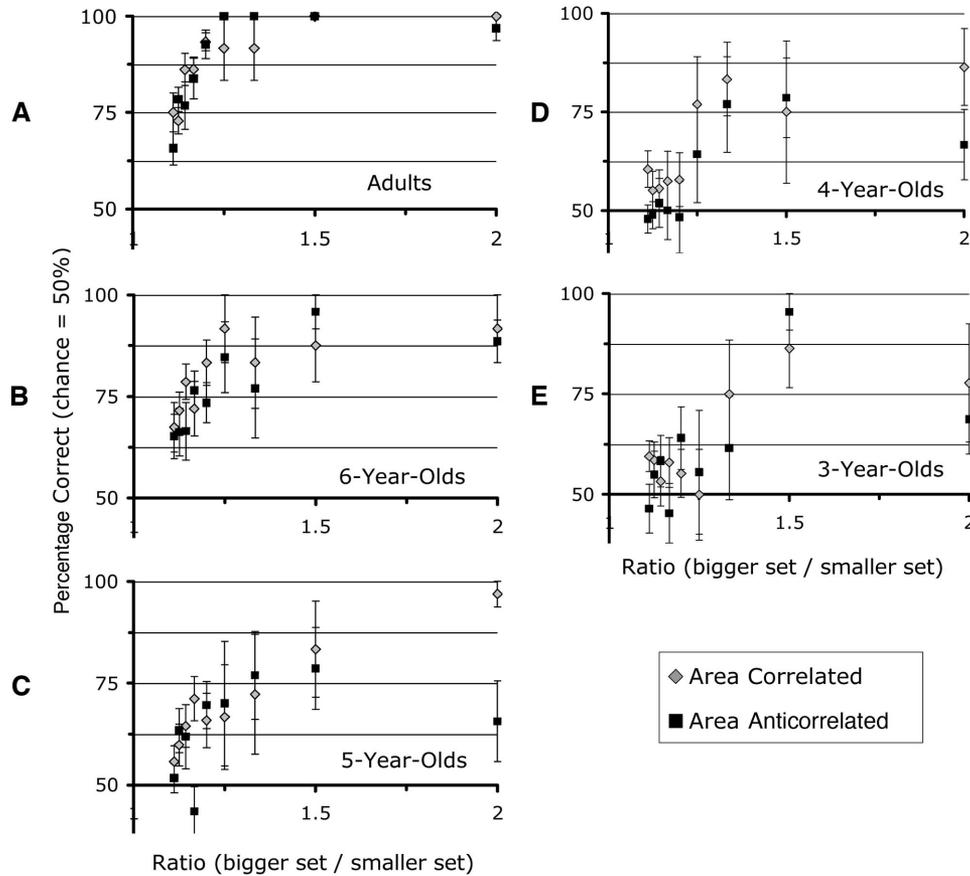


Figure 2. Scatterplots of percentage correct ($\pm SE$) on the Who Has More task as a function of ratio and trial type for each group (chance = 50%). A: adults, B: 6-year-olds, C: 5-year-olds, D: 4-year-olds, E: 3-year-olds.

the task. These data points artificially pull the performance of the psychophysics model down as it attempts to reduce the error between the fit and these data points, resulting in a Weber fraction that likely underestimates the actual acuity of these participants (witness the low R^2 values for these fits in Table 5).

To gain further accuracy in estimating the Weber fraction at each age (especially for younger children), we also modeled performance for each age group using a sigmoidal function fit by the Levenberg-Marquardt algorithm. We relied on the following sigmoidal equation:

$$y = \text{lower} + \frac{-\text{lower} + \text{upper}}{1 + e^{\frac{\text{inflection} - x}{\text{rate}}}} \quad (2)$$

This model should not be viewed as a psychologically plausible model of the representations involved, unlike the psychophysics model we use (Moyer & Bayer, 1976; Pica et al., 2004), but rather as a method for obtaining the best possible fit to our data in order to estimate the Weber fraction while controlling for participants' tendencies to guess randomly. The sigmoid model has a greater

Table 4
Percentage Correct Compared With Chance (50%) Across All Trials for Each Age Group

Age group	Area correlated				Area anticorrelated			
	Percentage correct		t	$p <$	Percentage correct		t	$p <$
M	SE	M			SE			
3 years	61.7	3.1	3.741	.005	60.1	3.2	3.178	.001
4 years	64.4	3.0	4.838	.001	58.5	2.6	3.256	.005
5 years	70.3	2.7	7.525	.001	63.7	3.8	3.596	.005
6 years	79.8	3.5	8.537	.001	76.1	2.0	13.280	.001
Adults	87.7	1.2	32.039	.001	87.2	1.2	31.099	.001

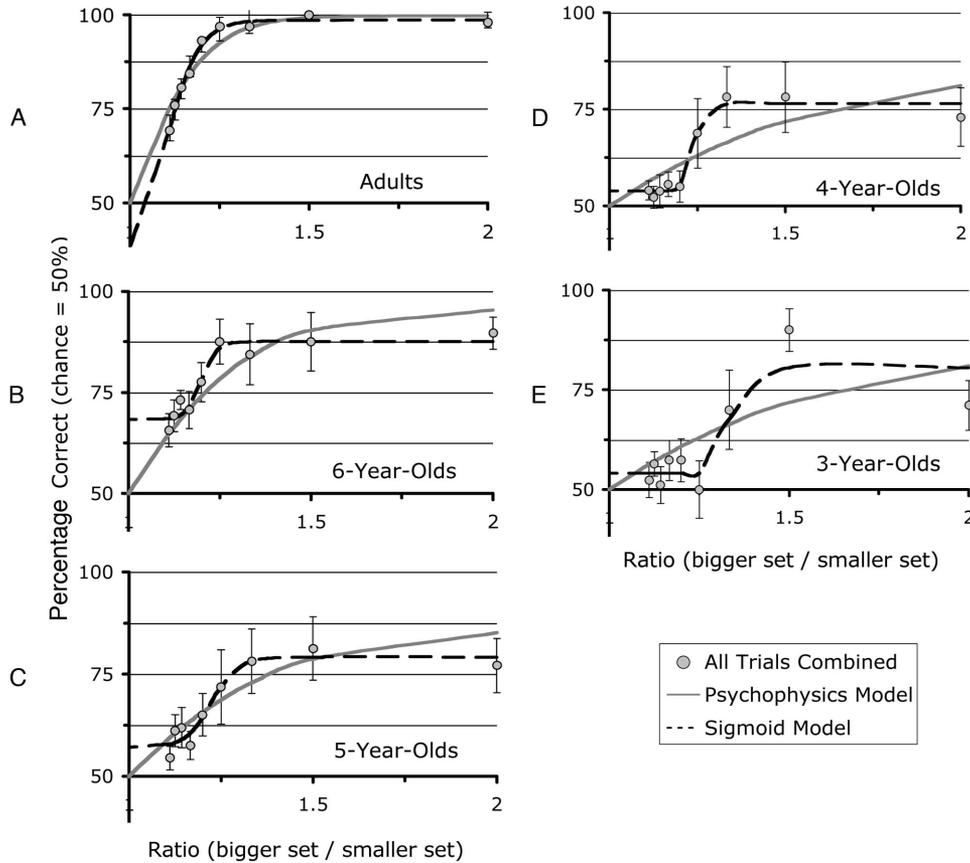


Figure 3. Scatterplots of percentage correct ($\pm SE$) on the Who Has More task, combined across trial type, as a function of ratio, presented with fits from a psychophysics model and a sigmoid model. A: adults, B: 6-year-olds, C: 5-year-olds, D: 4-year-olds, E: 3-year-olds.

number of free parameters than the psychophysics model and thus will nearly always provide a more accurate fit than the psychophysics model, but with reduced parsimony. From the sigmoidal equation for each age group, the Weber fraction can be estimated as the inflection point of the sigmoid (psychologically equivalent to the midpoint between successful discrimination [upper asymptote] and subjective-equality [lower asymptote]). As seen in Figures 3d and 3e and Table 5, the sigmoidal function provides an accurate fit to our data from younger children (e.g., 4-year-olds R^2

value = .98). Table 5 lists the estimated Weber fraction for each age group from both the psychophysics model and the sigmoid model along with the R^2 values for the fit. These numbers are also translated into a nearest whole number ratio (e.g., Weber fraction of .25 = 4:5 ratio). As Table 5 shows, estimated Weber fraction decreased with age, confirming that ANS acuity increased across the age groups we tested. Whereas 3-year-old children can accurately discriminate numerosities differing by a 3:4 ratio, 6-year-olds have sufficient acuity to discriminate numer-

Table 5
Estimated Weber Fraction (w)

Age group	Psychophysics model			Sigmoid model		
	w	R^2	Nearest whole number fraction	w	R^2	Nearest whole number fraction
3 years	.525	.502	2:3	.333	.829	3:4
4 years	.383	.632	3:4	.240	.976	4:5
5 years	.229	.785	4:5	.225	.938	5:6
6 years	.179	.846	6:7	.199	.940	5:6
Adults	.108	.926	9:10	.097	.988	10:11

Note. R^2 values represent the agreement between the modeled fit and the data for the entire function (see Figure 3).

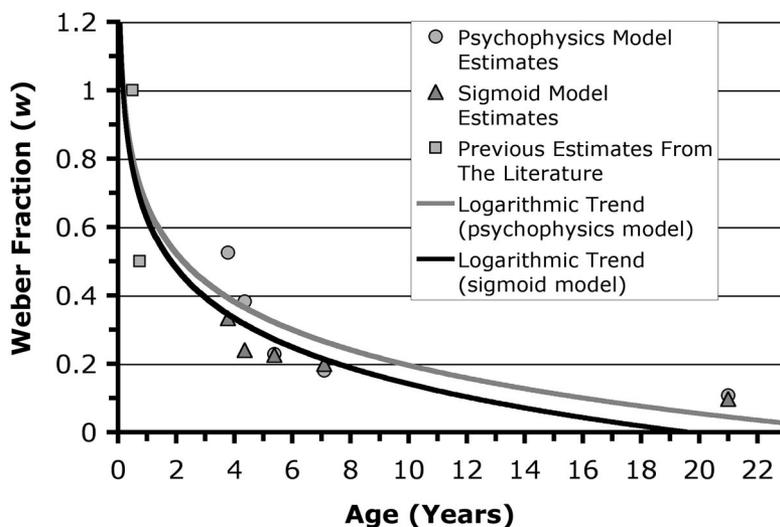


Figure 4. Scatterplot of the estimated Weber fraction for each age group from both the psychophysics model and the sigmoid model, combined with two estimates from the literature on infant Weber fraction displayed as a function of age. The estimated developmental curves are logarithmic least squares fits.

osities differing by a 5:6 ratio, and adults' acuity in our sample was as high as 10:11. Using a numerical discrimination task similar to ours, Pica et al., (2004) estimated untrained French adults' Weber fraction to be .12 or a 9:10 ratio. The extent of individual differences in Weber fraction in adults and children remains to be determined, but our estimate and that of Pica et al., suggest that the average acuity in adults from educated numerate cultures is in the range of 9:10 or 10:11.

In Figure 4 we have plotted estimates from both models alongside estimates from the developmental literature on infants' numerical acuity (Lipton & Spelke, 2003; Xu & Spelke, 2000). These estimates have been modeled by least squares fit to determine the developmental trajectory of the increasing acuity of the ANS (i.e., decreasing Weber fraction). These suggest a logarithmic decrease in Weber fraction throughout childhood,⁵ with adultlike levels of numerical acuity being attained sometime during the preteen years.

Discussion

Although children between 3 and 6 years old have already begun formal instruction in mathematics, the present results show that the acuity of the ANS is still developing during this time. Indeed, the sharpening of the ANS does not appear to be complete until early adolescence. Given the central role this system plays in supporting mathematical intuitions, this protracted period of development highlights the importance of coming to understand the effects of changes in ANS acuity on math learning and achievement (Booth & Siegler, 2006; Jordan et al., 2007).

What causes the ANS to increase in acuity? Although some of the sharpening of this system may be due to simple maturation of the neural circuitry subserving the ANS, recent evidence suggests that experience can also affect its development. Practice at numerical discrimination appears to increase acuity in children with math learning disabilities (Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006). After approximately 8 hr of practice spanning 5

weeks on a computer game designed to engage the ANS, 7- to 9-year-old children showed improvement in both symbolic and nonsymbolic numerical tasks. Continued engagement in numerical discrimination throughout childhood may therefore contribute to the increase in acuity we observed in normally developing children. This suggests the intriguing possibility that subpopulations that do not frequently engage in numerical discrimination might show increasingly reduced acuity with age relative to individuals undergoing formal education. Existing data are consistent with this hypothesis. Although Pica et al., (2004) stressed the similarity in the numerical discrimination abilities of their French and Mundurucu participants, Mundurucu adults' Weber fraction was estimated to be .17 (approximately 6:7), whereas French adults' was estimated to be .12 (approximately 9:10). This divergence could plausibly be the result of differences in day-to-day engagement in numerical discrimination. Training studies in normally developing children and adults will be helpful in determining whether practice with numerical discrimination leads to increases in acuity.

Changes in working memory capacity and executive functioning may also affect numerical discrimination. There are improvements in both spatial and verbal working memory in the preschool and early elementary school years (for a review, see Cowan, 1997), as

⁵ We have included logarithmic fits (in fact, negative ln fits) because these are perhaps more familiar to some readers. However, the data are more accurately fit by a power function with an exponent of $-.55$ (e.g., for previous data points from the literature and points from the current sigmoid model the least squares power function fit is $14.871x^{-.5535}$, $R^2 = .9488$). The power function has the psychologically sensible behavior of predicting asymptotic performance to be a Weber fraction > 0 (as opposed to a negative ln function that would allow for psychologically meaningless negative values for the Weber fraction as seen in the trend lines of Figure 4). The power function is the most likely group trend given the current data and should be used in all future studies. The logarithmic trends in Figure 4 are given solely for ease of communication with readers more familiar with these functions.

well as in aspects of executive functioning such as inhibition and cognitive flexibility (Espy, Kaufmann, McDiarmid, & Glisky, 1999; Happaney, Zelazo, & Stuss, 2004). These factors correlate with math achievement in tasks that focus on symbolic and explicit mathematical reasoning (Espy et al., 2004; Gathercole & Pickering, 2000; McClelland, Acock, & Morrison, 2006; McClelland et al., 2007). In our task, although working memory and other executive functions were likely needed to maintain multiple arrays in memory and to focus on the dimension relevant to the task (i.e., number rather than area), differences in these factors were minimized by lengthening the display times for younger children. Nonetheless, an important future direction is to determine how developmental changes in these nonnumerical abilities affect development of the ANS.

Orthogonal to changes in ANS acuity, we also found that 3- and 4-year-old children exhibited an increased likelihood of guessing randomly, especially on harder numerical comparisons. Because we presented the identical task and identical numerical ratios to all of our participants, task difficulty varied considerably with age. For adults, most of the numerical discriminations we presented were easy. For 3-year-olds, most of the discriminations were hard. It is possible that the high rates of guessing shown by the youngest children resulted from general discouragement in the face of so many hard problems. Developmental change in executive function is another likely contributor to this difference across age groups (Happaney et al., 2004). Six-year-olds remained vigilant and attempted to respond correctly on the most difficult discrimination trials (such as 9:10), even though the acuity of their ANS supported a level of accuracy of only 62% correct. That they remained vigilant is seen in Figure 3, where percentage correct on difficult ratios for this age group did not deviate from the predictions of the psychophysics model. Future studies may rely on a staircase approach that tailors the ratios presented to focus on a level of difficulty appropriate for each child (e.g., between 1:2 and 5:6 for 3-year-olds). Such an approach may result in lower guessing rates for younger children. It may not, however, reveal considerably different estimates in ANS acuity from those reported here, because the sigmoid model used here allowed us to control for rates of random guessing while estimating acuity. Still, the present cross-sectional data provide developmental milestones of ANS acuity that can guide future investigations into nonnumerical contributors to children's performance in numerical tasks.

Our task was designed to neutralize nonnumerical correlates of number by varying the degree to which number and both area and contour length were correlated across trial types. However, it is important to note that we used only a single task to estimate numerical acuity, and we did not measure any potential increases in acuity for any nonnumerical dimensions (e.g., changes in the ability to discriminate differences in surface area). Without such measures, it remains possible that the developmental changes identified in the present experiment are the result of more general developmental changes in magnitude representations, as opposed to changes that are specific to numerical representations. Previous research with both children and nonhuman animals suggests that the representation of numerical magnitude may rely on a format also shared by temporal magnitude (Brannon, Suanda, & Libertus, 2007; Feigenson, 2007; VanMarle & Wynn, 2006; Walsh, 2003) and perhaps by continuous extent (Brannon, Lutz, & Cordes, 2006; Feigenson, 2007; Walsh, 2003; Zorzi, Priftis, & Umiltà, 2002).

Whether number, time, and space all share a common representational format remains a question of great theoretical interest. The developmental changes in number discrimination that we describe here provide what we hope will be a useful comparison point for future work investigating developmental changes in the discrimination of time, area, volume, and other magnitudes.

In summary, the ability to nonverbally approximate number plays a role in quantitative reasoning throughout the human life span, even after the ability to represent exact integers is attained. Cognitive psychology and cognitive neuroscience have made much recent progress in understanding this ability. However, little research has addressed the development of the ANS after infancy. Here we show that the acuity of the ANS continues to increase between ages 3 and 6 years and does not reach adultlike levels until some time during the preteen years. The protracted nature of ANS development, spanning the period when symbolic mathematical instruction begins, has implications both for math education and for our understanding of the interplay between individual experience and the "number sense."

References

- Barth, H., Kanwisher, N., & Spelke, E. S. (2003). The construction of large number representations in adults. *Cognition*, *86*, 201–221.
- Barth, H., LaMont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. S. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222.
- Barth, H., LaMont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*, 14116–14121.
- Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *41*, 189–201.
- Brannon, E. M., Lutz, D., & Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Developmental Science*, *9*, F59–F64.
- Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental Science*, *10*, 770–777.
- Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.
- Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408–411.
- Cordes, S., Gelman, R., & Gallistel, C. R. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review*, *8*, 698–707.
- Cowan, N. (1997). The development of working memory. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 163–199). Hove, England: Psychology Press.
- Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 466–503). London: Oxford University Press.
- Droit-Volet, S., Clément, A., & Fayol, M. (2003). Time and number discrimination in a bisection task with a sequence of stimuli: A developmental approach. *Journal of Experimental Child Psychology*, *84*, 63–76.
- Espy, K. A., Kaufmann, P. M., McDiarmid, M. D., & Glisky, M. L. (1999).

- Executive functioning in preschool children: Performance on A-not-B and other delayed response format tasks. *Brain and Cognition*, 41, 178–199.
- Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematical skills in preschool children. *Developmental Neuropsychology*, 26, 465–486.
- Feigenson, L. (2007). The equality of quantity. *Trends in Cognitive Sciences*, 11(5), 185–187.
- Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, 44, 33–66.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307–314.
- Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, 44, 43–74.
- Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, 4, 59–65.
- Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology*, 70, 177–194.
- Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, 447, 589–591.
- Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, 17, 171–183.
- Green, D., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Krieger Publishing Company.
- Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science*, 17, 572–576.
- Happaney, K., Zelazo, P. D., & Stuss, D. T. (2004). Development of orbitofrontal function: Current themes and future directions. *Brain and Cognition*, 55, 1–10.
- Held, R., Birch, E., & Gwiazda, J. (1980). Stereoacuity of human infants. *Proceedings of the National Academy of Sciences*, 77, 5572–5574.
- Huntley-Fenner, G. (2001). Children's understanding of number is similar to adults' and rats': Numerical estimation by 5–7 year olds. *Cognition*, 78, B27–B40.
- Huntley-Fenner, G., & Cannon, E. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog magnitude mechanism. *Psychological Science*, 11, 147–152.
- Johnson, D. C. (1979). Teaching estimation and reasonableness of results. *Arithmetic Teacher*, 27, 34–35.
- Jordan, K. E., & Brannon, E. M. (2006). A common representational system governed by Weber's law: Nonverbal numerical similarity judgments in 6-year-olds and rhesus macaques. *Journal of Experimental Child Psychology*, 95, 215–229.
- Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research and Practice*, 23, 36–46.
- Levin, J. A. (1981). Estimation technique for arithmetic: Every day math and mathematics instruction. *Educational Studies in Mathematics*, 12, 421–431.
- Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large number discrimination in human infants. *Psychological Science*, 14, 396–401.
- Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, 5, 271–290.
- McClelland, M. M., Acocck, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, 21, 471–490.
- McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology*, 43, 947–959.
- McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, 15, 776–781.
- Moyer, R. S., & Bayer, R. H. (1976). Mental comparison and the symbolic distance effect. *Cognitive Psychology*, 8, 228–246.
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, 215, 1519–1520.
- Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences*, 101, 7457–7462.
- Piazza, M., Pinel, P., LeBihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, 53, 293–305.
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 306, 499–503.
- Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequalities. *Child Development*, 48, 630–633.
- Starkey, P., & Cooper, R. G. (1995). The development of subitizing in young children. *British Journal of Developmental Psychology*, 13, 399–420.
- Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences*, 95, 7836–7841.
- VanMarle, K., & Wynn, K. (2006). Six-month-old infants use analog magnitudes to represent duration. *Developmental Science*, 9, F41–F49.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space, and quantity. *Trends in Cognitive Sciences*, 7, 483–488.
- Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Non-verbal counting in humans: The psychophysics of number representation. *Psychological Science*, 10, 130–137.
- Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of "The Number Race," an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, 2, 20.
- Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science*, 8, 173–181.
- Wynn, K. (1992). Children's acquisition of number words and the counting system. *Cognitive Psychology*, 24, 220–251.
- Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, 89, B15–B25.
- Xu, F., & Arriaga, R. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology*, 25, 103–108.
- Xu, F., & Spelke, E. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74, B1–B11.
- Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, 8, 88–101.
- Zelazo, P. D., Craik, F. I. M., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica*, 115, 167–183.
- Zorzi, M., & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society* (pp. 778–783). Mahwah, NJ: Erlbaum.
- Zorzi, M., Priftis, K., & Umiltà, C. (2002). Brain damage: Neglect disrupts the mental number line. *Nature*, 417, 138–139.

Received October 1, 2007

Revision received April 3, 2008

Accepted April 8, 2008 ■